PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :
C12Q 1/68, 1/42, C12N 9/12
C12N 15/11, 15/66, C12P 19/34

A1

(11) International Publication Number: **WO 91/12341**

(43) International Publication Date: 22 August 1991 (22.08.91)

(21) International Application Number: PCT/US91/00951

(22) International Filing Date: 11 February 1991 (11.02.91)

(30) Priority data:
478,584 12 February 1990 (12.02.90) US

(60) Parent Application or Grant
(63) Related by Continuation
US 478,584 (CON)
Filed on 12 February 1990 (12.02.90)

(71) Applicant *(for all designated States except US)*: FRED HUTCHINSON CANCER RESEARCH CENTER [US/US]; 1124 Columbia Street, Seattle, WA 98104 (US).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)* : HENIKOFF, Steven [US/US]; 4711 - 51st Place SW, Seattle, WA 98116 (US).

(74) Agent: BRODERICK, Thomas, F.; Christensen, O'Connor, Johnson & Kindness, 2800 Pacific First Centre, 1420 Fifth Avenue, Seattle, WA 98101 (US).

(81) Designated States: AT (European patent), BE (European patent), CA, CH (European patent), DE (European patent), DK (European patent), ES (European patent), FR (European patent), GB (European patent), GR (European patent), IT (European patent), JP, LU (European patent), NL (European patent), SE (European patent), US.

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: IMPROVED METHOD FOR ORDERED DELETIONS FROM A FIXED POINT IN A CLONED INSERT

(57) Abstract

An improved method for generating ordered deletions from any fixed point in a cloned insert. Starting with a single-stranded phagemid template, DNA polymerase is used to extend an annealed primer. For DNA sequencing, the polymerase primer is typically selected to hybridize to the vector molecule near the 5′ side of the cloned insert. For *in vitro* mutagenesis from any point in the cloned insert, a custom polymerase primer is constructed to hybridize to the cloned insert just 5′ to the desired deletion site. Extention of the primer leads to a fully double-stranded circular molecule with a nick or small gap just 5′ to the primer. Exonuclease III initiates progressive digestion from the resulting 3′ end. Removal of times aliquots and digestion with a single-strand-specific endonuclease leads to a series of linear nested fragments of the insert, having a common end corresponding to the 5′ end of the primer. These molecules contain all of the vector DNA and so are readily circularized and used to transform cells, providing large numbers of deletion clones with targeted breakpoints.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**ES**	Spain	**MG**	Madagascar
AU	Australia	**FI**	Finland	**ML**	Mali
BB	Barbados	**FR**	France	**MN**	Mongolia
BE	Belgium	**GA**	Gabon	**MR**	Mauritania
BF	Burkina Faso	**GB**	United Kingdom	**MW**	Malawi
BG	Bulgaria	**GN**	Guinea	**NL**	Netherlands
BJ	Benin	**GR**	Greece	**NO**	Norway
BR	Brazil	**HU**	Hungary	**PL**	Poland
CA	Canada	**IT**	Italy	**RO**	Romania
CF	Central African Republic	**JP**	Japan	**SD**	Sudan
CG	Congo	**KP**	Democratic People's Republic of Korea	**SE**	Sweden
CH	Switzerland			**SN**	Senegal
CI	Côte d'Ivoire	**KR**	Republic of Korea	**SU**	Soviet Union
CM	Cameroon	**LI**	Liechtenstein	**TD**	Chad
CS	Czechoslovakia	**LK**	Sri Lanka	**TG**	Togo
DE	Germany	**LU**	Luxembourg	**US**	United States of America
DK	Denmark	**MC**	Monaco		

Improved method for ordered deletions from a fixed point in a Cloned Insert.

This is a continuation-in-part of pending prior application Serial No. 07/434,420, filed November 9, 1989, which is a continuation of prior application Serial No. 07/037,049, filed April 10, 1987 (now Patent No. 4,889,799), which is a divisional of application Serial No. 06/581,311, filed February 17, 1984 (now Patent No. 4,843,003). The aforesaid U.S. Patents No. 4,889,799 and 4,843,003 are hereby incorporated by reference.

This invention was made with government support under Grant No. GM29009 awarded by the National Institutes of Health. The government has certain rights in this invention.

Technical Field

This invention relates to genetic engineering, and particularly to the sequencing and/or functional analysis of genomic DNA.

Background of the Invention

Ordered deletion strategies have become popular both for DNA sequencing and for *in vitro* mutagenesis. Several of these strategies involve the random generation of deletions, either *in vitro* (Frischauf et al., 1980; Barnes & Bevan, 1983; Hong, 1982; Misra, 1987) or *in vivo* (Ahmed, 1985), followed by size selection on agarose gels. (For full citations, see the appended Literature Citations section.) Other methods take advantage of the properties of enzymes, usually exonucleases, which can generate deletions progressively, so that size selection is not required to obtain deletion breakpoints throughout a cloned segment (Poncz et al., 1982; Henikoff, 1984; Dale et al., 1985). Some of these latter methods involve unidirectional exonucleolytic digestion, using either *Escherichia coli* Exonuclease III (Henikoff, 1984; Yanisch-Perron et al., 1985; Ozhaynak and Putney, 1987) or the exonuclease activity of bacteriophage T_4 DNA polymerase

(Dale et al., 1985; Barcak & Wolf, 1986), so that circularization, rather than recloning of the treated molecules, yields targeted deletion clones with high efficiency. These unidirectional deletion methods are currently the most widely used.

Each of these methods has limitations, particularly with large cloned inserts. In the case of unidirectional deletions generated by Exonuclease III, two unique restriction enzyme sites are required in the polylinker region, one to block exonucleolytic digestion and one for accessibility to the enzyme. In addition, the variability in the purity of restriction enzymes and in the quality of the starting double-stranded template can lead to variable results, presumably because of nicks, which are attacked by Exonuclease III (Barnes & Bevan, 1983), or because of incomplete restriction enzyme digestion (Henikoff, 1987a; Nakayama & Nakauchi, 1989; Steggles, 1989). In the case of T_4 DNA polymerase digestion of a linearized single-stranded molecule linearized with a restriction enzyme, the polymerase does not digest at nearly as uniform a rate as Exonuclease III, leading to a broad range, rather than a tight clustering of targeted breakpoints (Dale et al., 1985). It also has been pointed out that the polymerase does not digest through certain sequences (Burton et al., 1988). A solution to this latter problem is to use polymerase extension of a primer annealed to a single-stranded template followed by single-strand-specific endonuclease treatment to generate fragments which are cloned into a double-stranded vector (Burton et al., 1988; Liu & Hackett, 1989). However, these methods are not very effective at targeting deletion breakpoints for larger inserts and furthermore require at least one unique restriction site.

Finally, all of the current *in vitro* methods generate nested deletions from an endonuclease cleavage site. This can severely limit the general utility of these strategies for constructing deletions to be used in functional analyses, since appropriate sites rarely coincide precisely with ends of functional regions.

Summary of the Invention

An improved method is described for generating ordered deletions from any fixed point in a cloned insert. Starting with a single-stranded phagemid template, DNA polymerase is used to extend an annealed primer. For DNA sequencing, the polymerase primer is typically selected to hybridize to the vector molecule near the 5' side of the cloned insert. For *in vitro* mutagenesis from any point in the cloned insert, a custom polymerase primer is constructed to hybridize to the cloned insert just 5' to the desired deletion site. In either embodiment, extension of the primer leads to a fully double-stranded circular molecule with a nick or small gap just 5' to the primer. Exonuclease III initiates progressive digestion

from the resulting 3' end. Removal of timed aliquots and digestion with a single-strand-specific endonuclease leads to a series of linear nested fragments of the insert, having a common end corresponding to the 5' end of the primer. These molecules contain all of the vector DNA and so are readily circularized and used to transform cells, providing large numbers of deletion clones with targeted breakpoints. This six-step procedure conveniently involves successive additions to tubes, beginning with a single-stranded template and ending with transformation—no extractions, precipitations, or centrifugations are needed. Results are comparable to those obtained with standard Exonuclease III-generated deletion protocols, but there is no requirement for restriction enzyme digestion or for highly purified double-stranded DNA starting material. This method is useful for obtaining nested deletions in either or both directions using a single parent clone, both for DNA sequencing and for functional analysis.

Reagent kits are also provided for practicing this improved method. Such kits typically provide a circular single-stranded recombinant vector molecule that includes a polylinker region having a first polymerase primer binding site 5' to a cloning region that has restriction endonuclease sites capable of receiving a target DNA segment. Such a vector molecule is typically provided in combination with (but packaged separately from) the following reagents: a first polymerase primer that is complementary to the first polymerase primer binding site in the vector molecule; a DNA polymerase that is capable of extending the first polymerase primer through the sequencing primer binding site and then the cloning region up to the 5' side of the first polymerase primer; an Exonuclease III reagent; and a single-strand-specific endonuclease. The kit may also include a DNA ligase, T_4 polynucleotide kinase, and one or more buffer solutions for the various reagents that are provided. Printed instructions are typically supplied with the reagent kit, describing how to practice the subject method using the reagents supplied in the kit. For greater versatility, the polylinker region of the vector molecule may have first and second restriction site sequences between the cloning region and a sequencing primer binding site. These first and second restriction sites are arranged so that cleavage of the polylinker region at the first and second restriction sites, following conversion of the circular single-stranded recombinant vector molecule into a circular double-stranded vector molecule, will produce a linearized doubl -stranded vector molecule having a first terminus adjacent the cloning region that is susceptible to digestion by Exonuclease III and a second terminus, not susceptible to digestion by Exonuclease III, adjacent the sequencing primer binding site.

Brief Description of the Drawings

FIGURE 1 presents an overview of the new method (steps 1-6), wherein for illustrative purposes the open box represents the primer binding site used for extension and also the fixed breakpoint for nested deletions in the resulting clones;

FIGURE 2 shows an application of the new method to a typical insert, by way of an agarose gel analysis of a deletion series (steps 1-4) carried out on a cDNA insert from the *D. melanogaster bw* locus cloned into pVZ1;

FIGURE 3 shows results of an Exonuclease III digestion series carried out as described for FIGURE 2, except that 2 µg single-stranded template was used, predigestion was done with *Msp*I, extension was carried out using cloned T_4 DNA polymerase, and single-stranded nuclease digestion using S1 nuclease;

FIGURE 4 shows an application of the new method to a large insert, by way of an agarose gel analysis of a T_4 DNA polymerase extension and Exonuclease III digestion of 4 µg single-stranded phagemid DNA generated from a 6201-bp *Sal*I genomic fragment from the *Drosophila pseudoobscura Gart* locus cloned into pVZ1;

FIGURE 5 shows mapping of 33 deletion breakpoints from a single early time point for a 6.2-kb cloned insert; and,

FIGURE 6 presents a vector (panel a) and strategy (panel b) for obtaining nested deletions for sequencing both strands of a single cloned insert.

Detailed Description of the Preferred Embodiment

Below, I describe a new method for the generation of ordered deletions that combines the advantages of several unidirectional deletion strategies. The method uses Exonuclease III, leading to a tight clustering of deletion breakpoints; however, in common with the T_4 DNA polymerase digestion strategy of Dale et al. (1985), this method uses single-stranded circular phagemid DNA (Mead & Kemper, 1986) as starting material and does not require unique restriction sites. Full extension of an oligonucleotide primer around the circle using T_4 DNA polymerase provides a double-stranded circular molecule with a nick (or small gap) adjacent to the 5' end of the primer. This nick is attacked by Exonuclease III, leading to the generation of targeted deletions. The six-step procedure described below requires relatively little labor and expense, as it involves successive additions to small microcentrifuge tubes through the transformation step. Each step can be conveniently monitored by agarose gel electrophoresis. By way of example, the new method is demonstrated below using inserts that range from 2.7-6.3 kb in length, with results comparable to those obtained using the now-standard

Exonuclease III digestion method (Henikoff, 1984) that is described in U.S. Patent No. 4,843,003, which is incorporated by reference herein. In addition to its usefulness in DNA sequencing, the improved method makes it possible to generate a nested set of targeted deletions for *in vitro* mutagenesis from any point in a cloned insert using a single custom primer.

A representative embodiment of the new method is outlined in FIGURE 1. Starting with single-stranded DNA obtained from helper-infected cells harboring a ColE1-derived phagemid, such as pKUN (Konings et al., 1987), pVZ1 (Henikoff & Eghtedarzadeh, 1987), Bluescript™ (Stratagene), or pGEM5Z (Promega), a fully double-stranded molecule is generated by polymerase extension of a primer. The primer (typically a 20-mer with a 3'-OH end) is complementary to a sequence (open box, step 1) within the vector adjacent to the insert, such that extension proceeds through the vector first. This leaves a nick or a very small gap (open box, step 2) adjacent to the 5' end of the primer. Addition of Exonuclease III leads to degradation of the synthetic strand starting at the 3' end of the nicked strand. At regular intervals, the uniformly digested gapped molecules are transferred to a solution containing a single-strand-specific endonuclease, such as S1 nuclease, and incubated to remove the single strand opposite to the gap. The ends are prepared for ligation using T_4 DNA polymerase and T_4 polynucleotide kinase (for the nonphosphorylated primer used in the extension) followed by ligation. Transformation of competent recA$^-$ *E. coli* cells leads to large numbers of clones with targeted breakpoints, most of which correspond to the extent of Exonuclease III digestion, as judged by agarose gel analysis of treated molecules. Transformants can be used to prepare double-stranded templates for dideoxy sequencing or for introduction into a host organism for functional analysis.

Details for carrying out this procedure are provided in the following Materials and Methods, which respectively describe preparation of exemplary solutions and a representative step-by-step procedure. The steps are: 1) annealing of the primer, optionally preceded by treatment of the template with a restriction enzyme to remove possible contaminating double-stranded molecules; 2) extension of the primer using T_4 DNA polymerase; 3) time-dependent gapping of the synthetic strand using Exonuclease III; 4) removal of single strands using S1 nuclease; 5) polishing and ligation using a mixture of T_4 DNA polymerase, T_4 polynucleotide kinase, and T_4 DNA ligase; and 6) transformation of competent host cells. All six steps (FIGURE 1) may be carried out by successive additions to 0.5 ml microfuge tubes and heat inactivations. No extractions, precipitations, or centrifugations are required.

Materials: Preparation of solutions and reagents

Stock solutions:

- *Template*: single-stranded phagemid DNA at 0.2-2 µg/µL H_2O or a low EDTA Tris buffer.
- *Primer*: αT3 20-mer (5' CCCTTTAGTGAGGGTTAATT 3'), SEQ. ID. No. 1, reverse hybridization 17-mer (5' GAAACAGCTATGACCAT 3'), SEQ. ID. No. 2, or the equivalent at 4 pmol/µL.
- *10X TM*: 0.66M Tris-HCl pH8, 30mM $MgCl_2$.
- *DTT*: 0.1M dithiothreitol.
- *dNTPs*: 2.5mM each of the 4 deoxynucleoside triphosphates.
- *S1 buffer concentrate*: 2.5M NaCl, 0.3M potassium acetate pH 4.6, 10mM $ZnSO_4$, 50% glycerol.
- *10X ligation buffer*: 0.5M Tris-HCl pH7.6, 0.1M $MgCl_2$, 10mM ATP.
- *PEG*: 50% (w/v) PEG (polyethylene glycol 6000-8000).

Enzyme reagents:

- *Restriction enzyme*: (optional) frequent cutter that can be heat-inactivated and that leaves Exonuclease III-sensitive ends (such as *Msp*I), tested for the absence of single-stranded nuclease activity.
- *BSA*: 1mg/ml Bovine serum albumin (nuclease-free).
- *T_4 DNA polymerase*: 1-10 units/µL, either cloned or from T_4-infected cells.
- *Exonuclease III*: ~150-200 units/µL.
- *T_4 polynucleotide kinase*: 3-10 units/µL.
- *T_4 DNA ligase*: 2-10 units/µL.

Working solutions:

- *S1 mix*: 27µL S1 stock, 173µL H_2O, 60 units S1 or mung bean nuclease, where 1 unit causes 1µg of nucleic acid to become perchloric acid soluble in 1 minute at 37°C. Prepare fresh and store on ice until use.
- *S1 stop*: 150 mM Trizma Base (no HCl), 25mM EDTA.
- *Ligation cocktail*: 290µL H_2O, 50µL 10 X ligation buffer, 50µL PEG, 5 µL DTT, 5µL dNTPs, 1 unit T_4 DNA polymerase, 10 units T_4 polynucleotide kinase, 20 units T_4 DNA ligase. Add enzymes just before use.
- *Host cells*: Frozen aliquot (0.5-1ml) of Ca^{++}-treated $recA^-$ cells such as HB101 (Maniatis et al., 1982).

Methods: Constructing nested deletions

Step 1: All operations are carried out in 0.5-ml microfuge tubes. Mix in a volume of 22µL: 2-4µg single-stranded phagemid DNA, 4µL 10X TM, 4 pmole αT3

(or equivalent) oligonucleotide. [Optional: Add 1 unit of a heat-inactivatable 4-cutter restriction enzyme (such as *Msp*I) to partially digest contaminating double-stranded plasmid, and incubate 10 minutes at 37°C.] Heat to 75°C 5 minutes, then allow to cool slowly to 37°C over a period of 30-60 minutes. Evaporation and condensation can be minimized by placing a piece of insulating styrofoam over the tube in an aluminum tube-heating block. Remove a 2μL aliquot for subsequent gel analyses.

Step 2: Mix in a volume of 20μL: 2μL DTT, 4μL dNTPs, 4μL BSA, 5 units T_4 DNA polymerase. Prewarm to 37°C and add to the primed DNA at 37°C. Incubate 2-8 hours; extension can be monitored by removing 0.5-1μL aliquots and electrophoresing on an agarose gel. Inactivate polymerase by heating for 10 minutes at 70°C.

Step 3: Prepare tubes with 3μL S1 mix on ice, one for each desired time point. (At 37°C, the rate of Exonuclease III digestion is 400-500 bases/min, increasing by 10% per 1°C in the range of 30-40°C. Choose a temperature and a time interval to give the desired range of deletions. It is very important to maintain a uniform temperature within the reaction solution during the incubation to obtain tightly clustered deletions; it helps to use a tube-heating block as a preheated dispenser for the pipettor.) Warm to the desired temperature a tube containing enough polymerase-extended circle to provide 1μL per time point. Add 1/10 volume Exonuclease III, mixing thoroughly and rapidly with the pipettor. Remove successive timed 1μL aliquots to the S1 mix tubes, pipetting up and down to mix. Hold on ice until all aliquots are taken.

Step 4: Remove tubes to room temperature and incubate 15-30 minutes. Add 1μL S1 stop, removing a 1μL aliquot for gel analysis at the same time. Heat to 70°C for 10 minutes.

Step 5: Add 16μL ligation cocktail. Incubate the resulting ligation mix at room temperature 1 hour to overnight. (The number of transformants obtained increases with long incubations.)

Step 6: Transform by mixing 5μL ligation mix with 15μL freshly thawed Ca^{++}-treated cells on ice. After 30 minutes, heat-shock 1.5 minutes at 42°C, add 100μL SOC (Hanahan, 1985) or Luria broth and allow to recover about 30 minutes at 37°C before plating. Yields typically range from 10 to 1000 transformants depending on the ligation time and the competence of Ca^{++}-treated cells. The remaining ligation mixtures can be stored frozen for later use after overnight incubation.

Results

Application of the method to a typical insert: This procedure was carried out on a 2748 bp *Drosophila* cDNA insert (Dreesen et al., 1988) cloned into phagemid pVZ1. This vector has the bacteriophage T_3 RNA polymerase promoter adjacent to the polylinker into which the cDNA was inserted. Phagemid particles were produced by infection with M13KO7 under kanamycin and ampicillin selection (Vieira and Messing, 1987), and the resulting supernatants were used to prepare single-stranded circles by phenol-chloroform extractions of polyethylene glycol-precipitated particles (Anderson, 1981). The results are illustrated in FIGURE 2. A 20-mer complementary to the region of the T_3 promoter ("α-T_3": 5'-CCCTTTAGTGAGGGTTAATT-3') (SEQ. ID. No. 1) was annealed in 2-fold molar excess to the template (FIGURE 2, lane 1), and T_4 DNA polymerase was used to extend the primer. A sample electrophoresed on a 0.7% agarose gel after 30 minutes shows only partial extension (lane 2), whereas after 4 hours nearly all of the material migrates as expected for nicked circles (lane 3). Exonuclease III was added to this fully extended sample, and aliquots were removed at 30-second intervals. Alternative samples were treated with either S1 (lanes 5-10) or mung bean nuclease (lanes 12-16), and portions were electrophoresed. Exonuclease III digestion proceeded synchronously at a rate of about 400 bases per minute, as evidenced by the uniform size range of molecules at each time point for both S1 and mung bean nuclease treated samples.

Considering FIGURE 2 in greater detail, annealing to α-T_3 primer followed partial digestion of 4 μg single-stranded phagemid DNA with *Eco*RI (lane 1). Extension with T_4 DNA polymerase (Boehringer Mannheim, 1 unit/μL) was carried out according to the above-stated Methods, except that dNTPs were at a concentration of 0.15 mM. Aliquots were taken at 30 min (lane 2) and 4 hours (lane 3). Shortly thereafter the reaction was terminated by incubation at 70°C and placed at 37°C for digestion with Exonuclease III (Boehringer Mannheim, 175 units/μL). One-microliter aliquots were removed at 30-second intervals and placed alternately into 3μL S1 nuclease (Promega, 50 units/μL) or mung bean nuclease (Bethesda Research Labs, 50 units/μL) and digested 30 minutes at room temperature. S1 stop was added and the aliquots were loaded onto the gel. Each sample represents 1/40 of the starting material. Samples are: (lanes 5-10) 0', 1', 2', 3', 4', 5' Exonuclease III digestion followed by S1 nuclease treatment, and (lanes 12-16) 0.5', 1.5', 2.5', 3.5', 4.5' Exonuclease III digestion followed by mung bean nuclease treatment. Markers (lanes 4 and 11) are 1-kb ladder fragments with sizes (in bp from top to bottom) 12216, 11198, 10180, 9162, 8144, 7126, 6108,

5090, 4072, 3054, 2036, and 1636 (Bethesda Research Labs). The 0.7% agarose gel in TBE buffer (Maniatis et al., 1982) contained 0.2 µg/ml ethidium bromide.

In a similar experiment using this template, single-stranded circles were partially digested with *MspI* (GGCC, Bethesda Research Labs) in order to fragment supercoiled circles that occasionally contaminate template preparations, presumably due to cell lysis during infection. This was done in the annealing mixture and was followed by heat treatment, necessary both to inactivate the restriction endonuclease and to start the annealing reaction. After extension using T_4 DNA polymerase (U.S. Biochemical), Exonuclease III and S1 nuclease treatments, samples were mixed with ligation cocktail, incubated overnight, and 1/4 of the ligation mix was used to transform competent cells. This resulted in 200-400 ampicillin-resistant colonies from each time point. Five colonies from each time point were inoculated into 1.5 ml cultures. These were harvested for supercoil minipreps using a standard alkaline lysis procedure with a single phenol-chloroform extraction and ethanol precipitation (Chen & Seeburg, 1985). Supercoils from 70 clones were partially sequenced ("G-tracked") (Anderson, 1981) by the dideoxy method (Sanger et al., 1977) using a "reverse hybridization primer", which primes 38 bases upstream of the T_3 primer binding site. Sequencing was carried out using Sequenase™ (US Biochemical), ^{35}S dATP, and Mn^{++} substituted for Mg^{++} (to favor sequence close to the primer). The breakpoints for 59 clones from which sequence adequate for precise mapping was obtained are shown in FIGURE 3 as a function of the time of Exonuclease III digestion. The partial sequences of 6 clones were not sufficient for mapping either to the insert or the vector. Five clones did not yield any sequence, even though plasmid was present, suggesting loss of the primer binding site. Except for these 5 clones, the 5'-most base of the αT_3 primer was found adjacent to the breakpoint in each case.

Most of the breakpoints map very close to the dashed line representing a rate of 400 nucleotides per minute. With one exception, the other breakpoints are above the line, indicating inserts smaller than expected. These might have resulted from the occasional failure of the polymerase to fully extend, coupled with preferential ligation of the smaller circles. Alternatively, nonspecific priming within the insert region would lead to similar artifacts. For this reason, only a 2- to 4-fold excess of primer over template is annealed prior to extension. Four clones were not deleted at all. Interestingly, three of these were from the 1-minute time point (FIGURE 3). Although this clustering could have been coincidental, a more likely explanation is that, during transfer of this aliquot to the S1 mix, the pipettor touched the side of the tube above the solution, leaving a

small portion that escaped S1 digestion but not the subsequent end repair and ligation. To avoid this potential problem, the aliquots may be centrifuged briefly after removal from ice for S1 digestion.

Application of the method to a large insert: This method also has been used with larger templates. FIGURE 4 shows a time course for T_4 DNA polymerase extension of phagemid clones that has an insert of 6.2 kb, i.e., a 6201-bp *SalI* genomic fragment from the *Drosophila pseudoobscura Gart* locus (Henikoff & Eghtedarzadeh, 1987), cloned into pVZ1. Examination of aliquots electrophoresed on a 0.6% agarose gel reveals that extension proceeded slowly, and was not complete even after 6 hours (lanes 3-6). However, overnight extension was sufficient to complete the reaction, with better results when supplemented with increased polymerase and deoxynucleoside triphosphates (lanes 7-8). The appearance and eventual disappearance of bands during the reaction reflects pausing of the polymerase at specific sites during synthesis (Huang et al., 1981; Roth et al., 1982). A minor product which comigrated with full-length linear molecules is seen in each case (see also FIGURE 2, lane 4). This might have resulted from extension of full-length single-stranded linear molecules from "looped-back" 3' ends (Goulian et al., 1968). Most of the completely extended molecules were resistant to brief S1 nuclease treatment, as expected for nearly fully double-stranded molecules (FIGURE 4, lane 9). However, when S1 nuclease treatment was preceded by digestion with Exonuclease III, all of the molecules become susceptible to S1 nuclease attack (lane 10). In this case, 1.3 kb was removed, yielding molecules of about 8 kb.

Considering FIGURE 4 in more detail, the following apply. Lane 1: starting single-stranded material; lane 2: annealed to α-T_3 primer; lanes 3-7: T_4 DNA polymerase extensions for 1, 2, 3, 6, and 20 hours, respectively; lane 8: same as in lane 7 except that after 6 hours the reaction was supplemented by addition of 0.1 units/μL polymerase and 0.4 mM dNTPs; lane 9: S1 nuclease treatment of the sample electrophoresed in lane 8; lane 10: same, except after a 6-minute Exonuclease III digestion. Markers run in a nearby lane are indicated by arrows representing 9162 and 8194 bp. Aliquots represent 1/40 of the starting material. The 0.6% agarose gel in TBE buffer contained 0.2% ethidium bromide.

A portion of the sample shown in lane 10 of FIGURE 4 was used for ligation and transformation. The portion was incubated for 4 hours after addition of ligation cocktail and used to transform *E. coli*. The arrow in FIGURE 5 indicates the approximate extent of digestion based on the gel analysis in FIGURE 4. The distribution of deletion breakpoints for 36 of the resulting clones is shown in

FIGURE 5, except for 3 plasmids which were not interpretable by restriction mapping. Three plasmids were not detectably deleted (the optional restriction digestion of step 1 was omitted), 12 were of the desired size with breakpoints at about 1.3 kb into the insert, 13 were smaller with breakpoints within the 6.2 kb insert, and 5 were smaller with breakpoints within the vector. This result indicates that the method can be used effectively for targeting breakpoints in relatively large inserts.

T_4 DNA polymerase remains active even after several hours of digestion (FIGURE 4). One should be aware, however, that the highly active 3'-5' exonuclease activity will cause a "stuttering" of the polymerase after completion of extension; this might eventually deplete the reaction of deoxynucleoside triphosphates (dNTPs) leading to exonucleolytic degradation of the synthesized strand (Roth et al., 1982; data not shown). The ability of S1 nuclease to fully linearize the polymerase extended sample shown in FIGURE 2 (compare lanes 3 and 5) but not the sample shown in FIGURE 4 (compare lanes 8 and 9) is attributable to the lower dNTP concentration during the extension reaction in the former case, but not in the latter. It appears that successive depletion of as many as 3 of the 4 dNTPs will lead to a very small gap, which will have little consequence for this method. However, when all 4 dNTPs are depleted, the 3'-5' exonucleolytic activity can rapidly degrade the synthesized strand. For this reason, when incubating for extended periods, the reaction should be supplemented with more dNTPs. Concentrations as high as 0.5 mM in each dNTP have been used without apparent ill effects. Since some chelation of Mg^{++} occurs, an equimolar amount of $MgCl_2$ also should be added if higher levels of dNTPs are found to be necessary.

In the procedure described here, the primer used lacks a phosphate at its 5' end. For efficient ligation, this phosphate should be present. Since T_4 polynucleotide kinase requires the same buffer components as T_4 DNA ligase, it is convenient to carry out phosphorylation during ligation. Should a phosphorylated primer be used, it might be important to remove all traces of ATP before using the primer, since any T_4 DNA ligase contaminating a T_4 polymerase preparation from infected cells could ligate the nick during the long extension period.

Discussion

Comparison to the standard strategy for Exonuclease III generation of deletions: The method described here uses polymerase extension of a primed single-stranded circle to prepare a substrate for Exonuclease III digestion. One advantage of this approach over the now-standard Exonuclease III deletion strategy is that restriction enzyme digestions are not needed. In the latter case,

two unique restriction sites must be present in the polylinker region of the vector. The one closest to the primer binding site should have a 3' 4-base overhang and the other should not. These requirements have led to special vectors that usually overcome this problem, such as pVZ1 (Henikoff & Eghtedarzadeh, 1987), Bluescript™ (Stratagene), and pGEM5Z (Promega), although for larger inserts even these vectors are sometimes inadequate.

Probably a more serious practical problem is the variability in quality of restriction enzymes, especially since manufacturers do not test for nicking activity. As is evident from this work, and that of others (Taylor et al., 1985), Exonuclease III efficiently attacks the 3' end at a nick, so that only unnicked molecules in a preparation of restricted fragments will yield targeted deletion breakpoints. For the standard Exonuclease III procedure to be effective, digestion by both enzymes must be complete. Since the sites are so close, it is often difficult to determine this by agarose gel analysis. Thus, extended digestions are sometimes carried out, leading to increased nicking of the template, resulting in a preponderance of nontargeted deletions and even frequent loss of the primer binding site. One solution to these problems is to purify the successive aliquots after S1 digestion using low-melting-point agarose gels (Nakayama & Nakauchi, 1989; Steggles, 1989). While this step can lead to excellent results, it is relatively labor-intensive. Other problems are that restriction enzymes do not always digest well in a common buffer, and that phenol extraction and ethanol precipitation steps are generally necessary after digestion with restriction enzymes. It is noteworthy that the method described here avoids all of these problems, yet still yields comparable results. Furthermore, a single enzyme, T_4 DNA polymerase, provides a substrate for Exonuclease III digestion, greatly simplifying the quality control problems that arise when any of a multitude of restriction enzymes are needed. Another advantage is that the current method uses single-stranded templates, which are easily purified in high yields from the supernatants of infected cells.

A potential disadvantage of this method is that the resulting deletion clones cannot be sequenced using single-stranded templates produced by infection with the same helper phage that was used to generate the parent template. This is because the polarity of the primer used to extend the parent template is opposite to that needed for sequencing through the breakpoint of the resulting deletion template (see FIGURE 1). However, this can be avoided by the use of pKUN vectors, which have a second viral origin of replication derived from phage IKe oppositely oriented to the origin for phage M13 (Konings et al., 1987; Peeters

et al., 1986; Konings et al., 1986). Infection with a chimeric M13-IKe helper such as *Mike* yields pure single-stranded particles suitable for extension or sequencing. Even using vectors with a single viral origin, such as was done in the Examples described herein, rapid microfuge-tube-scale methods of preparing plasmid DNA supercoils for sequencing (Chen & Seeburg, 1985; Wang et al., 1988; Del Sal et al., 1989) can be used in conjunction with the current method (although clones in M13 phage vectors are less appropriate). It has been pointed out that some previous problems with supercoil sequencing stem from the use of $recA^+$ strains for the growth of plasmids (W. Salser, personal communication). It is thought that the concatemerized supercoils in these preparations give rise to tangled denatured molecules that have a much lower probability of any particular sequence being exposed to the solution. Therefore, much of the variability seen for double-stranded sequencing might be avoidable.

Choice of T_4 DNA polymerase for extension: Whereas the strong 3'-5' exonuclease activity might seem to be a disadvantage in the extension reaction, since it greatly reduces the rate of polymerization, it appears that this activity is essential to prevent strand displacement which occurs with other polymerases, such as PolIK (Klenow) and modified T_7 DNA polymerase (Lechner, 1983; data not shown). On the one hand, strand displacement activity allows an enzyme to melt out hairpins during polymerization; T_4 DNA polymerase is known to be impeded by such regions in substrate molecules (Huang et al., 1981; Roth et al., 1982; see FIGURE 4). Although this might be a problem in some cases, it appears that, given enough time, this enzyme can extend through most sequences, stopping when it reaches the 5' end of the primer. On the other hand, a polymerase that can cause strand displacement will continue beyond the 5' end of the primer by displacing this end, synthesizing a rolling circle. Such molecules are unsuitable as substrates for this method (data not shown).

Another feature of T_4 DNA polymerase that make it suitable for this method is the exceptionally high fidelity with which it replicates DNA. Its error frequency *in vitro* has been estimated to be no more than 10^{-7}, or about 10- to 100-fold lower than that for DNA polymerase I (Kunkel et al., 1984). This is a particularly important consideration for the generation of clones to be used in functional analyses and in DNA sequencing. The enzyme also completely lacks a 5'-3' exonuclease activity, so it will not nick-translate. a problem for this method similar to strand-displacement. Although it requires relatively long incubations to complete the reaction, the enzyme is sufficiently stable to make these incubations practical. Furthermore, both the cloned enzyme and that from T_4-infected cells

have yielded excellent results, even after very long incubations, indicating a lack of detectable nuclease activity in samples obtained from different commercial sources. The enzyme also is widely available and relatively inexpensive, as are the other components of this method.

Another DNA polymerase for use in the subject method is *E. coli* DNA polymerase III holoenzyme (Tsurushita et al., 1988; Nielson & Mathur, 1990).

Exemplary modifications and extensions of the method: The pivotal feature of this method is the preparation of a substrate for Exonuclease III. Subsequent steps can be modified as described in several previous studies. For example, problem templates that fail to extend quantitatively can be gel-purified after S1 or mung bean nuclease treatment as described (Nakayama & Nakauchi, 1989; Steggles, 1989). This might be useful for obtaining targeted deletions for early time points of large inserts, where the presence of even low levels of smaller molecules during the ligation can reduce the frequency with which targeted deletions are obtained (FIGURES 4 and 5). In addition, the Exonuclease III reaction can be slowed down in a controlled manner by reduction in temperature or enzyme concentration (Henikoff, 1987a) or by addition of NaCl (Tomb & Barcak, 1989). Moreover, when large numbers of aliquots are collected, V-bottom microtiter plates can be used (Henikoff, 1987b). The current method is readily automated, since it requires only successive additions and temperature changes.

The method described here can be applied to both strands of a single DNA segment by cloning it into a vector in both orientations. In addition, the procedure can supplement any of the widely used unidirectional deletion methods in providing a strategy to obtain sequence from both strands of a single clone. Using the standard Exonuclease III procedure, full sequence from both strands can be obtained only if 4 unique sites are available in appropriate positions on either side of the insert. A more generally applicable alternative is to use a vector such as pVZ1, which segregates the two classes of sites needed for the standard procedure to one side of the multiple cloning region (FIGURE 6, panel A). This allows deletions to be generated for one strand. The current method can be used to generate deletions for sequencing the opposite strand (FIGURE 6, panel B). Considering FIGURE 6 in more detail, panel A shows that the polylinker of pVZ1 separates sites for cloning (derived from M13mp18; Yanisch-Perron et al., 1985) from sites for accessibility (*Not*I or *Nar*I) and resistance (*Bbe*I, *Bst*I, *Nsi*I, or *Apa*I) to Exonuclease III in a phagemid vector. Panel B shows how a single-stranded template is used to generate a deletion series for double-stranded sequencing as described here, while the plasmid form is used to generate a standard deletion

series using restriction enzymes for single- or double-stranded sequencing (Henikoff, 1987a).

The strategy of Dale et al. (1985) is inherently constrained to provide sequence from only one strand. The current method also can supplement this approach by providing deletions for sequencing the opposite strand using the same single-stranded circular starting material. These mixed strategies are likely to be most useful in cases where obtaining both orientations of a single insert requires an extra cloning step. For example, in the selection of cDNA clones from libraries made in vectors such as Lambda ZAP™ (Stratagene), the resulting product is an insert already cloned into a phagemid vector.

In a preferred embodiment, the method described herein is applied to both strands of a single clone using a pKUN vector (Konings et al., 1987). Infection with an M13 or f1 helper such as M13K07 to obtain one strand for polymerase extension and with a *Mike* helper to obtain the other strand allows one to generate deletion series in either direction. Single-stranded templates for sequencing from the resulting clones can be obtained by infection with *Mike* in the first case and with M13K07 in the second. This strategy is likely to be particularly useful in cases where obtaining both orientations of a single insert requires an extra cloning step. For example, in the selection of cDNA clones from libraries made in vectors such as Lambda ZAP™ (Stratagene), the resulting product is an insert already cloned into a phagemid vector.

Previous methods for generating ordered deletions have been used primarily as subcloning strategies for DNA sequencing. For *in vitro* deletion mutagenesis, these methods have been limited by their requirement for a restriction site at the point from which deletions are generated. Since the present method has no such requirement, it allows for nested deletions to be made from any point in a cloned insert using a custom primer. Unlike current procedures for site-specific generation of deletions (Eghtedarzadeh & Henikoff, 1986; Chang et al., 1988; Wang et al., 1989), this method can be used for obtaining many different breakpoints from a single primer. These deletions are generated at nearly 100% efficiency, with a large fraction targeted to a predetermined region, making this method potentially applicable to many situations in which deletions are desirable. In addition, the use of the pKUN vector system (Konings et al., 1987) permits such deletions to be made in either direction from a single parent clone.

Very recently, another method for the generation of templates for dideoxy sequencing was introduced (Sorge & Blinderman, 1989). Successive aliquots from a timed digestion using Exonuclease III are used for dideoxy sequencing (Guo & Wu,

1982). The deoxycytosine triphosphates in this reaction are methylated, so that the synthetic bases are resistant to digestion by particular restriction enzymes. Restriction by one of these enzymes results in a set of labeled fragments having a unique end close to the point at which Exonuclease III stops. In this ExoMeth™ sequencing procedure (Stratagene; *Strategies in Molecular Biology* 3(1):14, 1990), successive aliquots from a standard Exonuclease III series are therefore used directly for sequencing rather than for constructing ordered deletion subclones. Depending on the distribution of sites for the restriction enzyme or enzymes used, most of the sequence from an insert can be obtained in this way. The method described here for synthesizing a substrate for Exonuclease III digestion provides an alternative method of preparing template for ExoMeth™ sequencing. Pursuant to this disclosure, sequences from both strands are advantageously obtained in this manner from a single clone using a pKUN vector (Konings et al., 1987).

Literature Citations

Ahmed, A., *Gene* 39:305-310, 1985.

Anderson, S., Shotgun DNA sequencing strategy using cloned DNase I-generated fragments, *Nucleic Acids Res.* 9:3015-3027, 1981.

Barcak, G.J., *Gene* 49:119-128, 1986.

Barnes, W.M., and M. Bevan, Kilo-sequencing: an ordered strategy for rapid DNA sequence data acquisition, *Nucleic Acids Res.* 11:349-368, 1983.

Vieira, J., and J. Messing, *Meth. Enzymol.* 153:3-11, 1987.

Burton, F.H. et al., A directed nucleotide-sequencing approach for single-stranded vectors based on recloning intermediates of a progressive DNA synthesis reaction, *Gene* 67:159-168, 1988.

Chang, G.J., et al., *DNA* 7:211-217, 1988.

Chen, E.Y., and P.A. Seeburg, *DNA* 4:165-170, 1985.

Dale, R.M.K., et al., A rapid single-stranded cloning strategy for producing a sequential series of overlapping clones for use in DNA sequencing: application to sequencing the corn mitochondrial 18S rDNA, *Plasmid* 13:31-40, 1985.

Del Sal, G., et al., *Biotechniques* 7:514-520, 1989.

Dreesen, T.D., et al., *Mol. and Cell Biol.* 8:5206-5215, 1988.

Eghtedarzadeh, M.K., and S. Henikoff, *Nucleic Acids Res.* 14:5115, 1986.

Frischauf, A.M., et al., A subcloning strategy for DNA sequence analysis, *Nucleic Acids Res.* 8:5541-5549, 1980.

Goulian, M., et al., *J. Biol. Chem.* 243:627-638, 1968.

Guo, L-H., and R. Wu, New rapid methods for DNA sequencing based on Exonuclease III digestion followed by repair synthesis, *Nucleic Acids Res.* 10:2065-2084, 1982.

Hanahan, D., in DNA Cloning: A practical approach, Vol. 1, D.M. Glover (ed.), IRL Press, Oxford, pp. 109-135, 1985.

Henikoff, S., Unidirectional digestion with Exonuclease III creates targeted breakpoints for DNA sequencing, *Gene* 28:351-359, 1984.

Henikoff, S., Unidirectional digestion with Exonuclease III in DNA sequence analysis, *Meth. Enzymol.* 155:156-165, 1987a.

Henikoff, S., Exonuclease III generated deletions for DNA sequence analysis, *Promega Notes* 8:1-3, 1987b.

Henikoff, S., and M.K. Eghtedarzadeh, Conserved arrangement of nested genes at the *Drosophila Gart* locus, *Genetics* 117:711-725, 1987.

Hong, G.F., A systematic DNA sequencing method, *J. Mol. Biol.* 158:539-549, 1982.

Huang, C.C., et al., *J. Biol. Chem.* **256**:4087-4094, 1981.

Konings, R.N.H., et al., *Gene* **46**:269-276, 1986.

Konings, R.N.H., et al., *Meth. Enzymol.* **153**:12-35, 1987.

Kunkel, T., et al., *J. Biol. Chem.* **259**:1539-1545, 1984.

Lechner, R.L., et al., *J. Biol. Chem.* **258**:11174-11184, 1983.

Liu, Z., and P.B. Hackett, *Biotechniques* **7**:722-728, 1989.

Maniatis, T., et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1982.

Mead, D.A., and B. Kemper, in Vectors: A survey of molecular cloning vectors and their uses, R.L. Rodriguez (ed.), Butterworth, Boston, 1986.

Misra, T.K., *Gene* **34**:263-268, 1987.

Nakayama, K., and H. Nakauchi, An improved method to make sequential deletion mutants for DNA sequencing, *Trends in Genetics* **5**:325, 1989.

Nielson, K., and E. Mathur, *Strategies in Molecular Biology* **3**(1):10, 1990.

Ozhaynak, E., and S.D. Putney, *Biotechniques* **5**:770-773, 1987.

Peeters, B.P.H., et al., *Gene* **41**:39-46, 1986.

Poncz, M., et al., "Nonrandom" DNA sequence analysis in bacteriophage M13 by the dideoxy chain-termination method, *Proc. Natl. Acad. Sci. USA* **79**:4298-4302, 1982.

Roth, A.C., et al., *J. Biol. Chem.* **257**:1267-1273, 1982.

Sanger, F., et al., DNA sequencing with chain-terminating inhibitors, *Proc Natl. Acad. Sci. USA* **74**:5463-5467, 1977.

Sorge, J.A., and L.A. Blinderman, ExoMeth sequencing of DNA: eliminating the need for subcloning and oligonucleotide primers, *Proc. Natl. Acad. Sci. USA* **86**:9208-9212, 1989.

Steggles, A.W., *Biotechniques* **7**:241-242, 1989.

Taylor, J.W., et al., *Nucleic Acids Res.* **13**:8765-8785, 1985.

Tomb, J.-F., and G.J. Barcak, *Biotechniques* **7**:932-933, 1989.

Tsurushita, N., et al., Site-directed mutagenesis with *Escherichia coli* DNA polymerase III holoenzyme, *Gene* **62**:135-139, 1988.

Vieira, J., and J. Messing, *Meth. Enzymol.* **153**:3-11, 1987.

Wang, L.-M., et al., *Biotechniques* **6**:839-843, 1988.

Wang, L.-M., et al., *Biotechniques* **7**:1000-1010, 1989.

Yanisch-Perron, C., et al., Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors, *Gene* **33**:103-119, 1985.

While preferred embodiments of the invention have been illustrated and described, it is to be understood that, within the scope of the appended claims, various changes can be made therein. Hence, the invention can be practiced in ways other than those specifically described herein.

The embodiments of the invention in which an exclusive property or privilege is claimed are defined as follows:

1. A reagent kit for producing shortened DNA fragments of a target DNA segment, comprising in combination:

a circular single-stranded recombinant vector molecule consisting essentially of vector DNA comprising a polylinker region, the polylinker region comprising a first polymerase primer binding site 5' to a cloning region, the cloning region comprising restriction endonuclease sites capable of receiving a target DNA segment;

a first polymerase primer complementary to the first polymerase primer binding site in the vector molecule;

a DNA polymerase capable of extending the first polymerase primer through first the sequencing primer binding site and then the cloning region up to the 5' side of the first polymerase primer;

Exonuclease III; and

a single-strand-specific endonuclease.

2. The reagent kit of Claim 1, wherein the polylinker region of the vector molecule comprises first and second restriction sites between the cloning region and a 3' sequencing primer binding site, the first and second restriction sites being disposed such that cleavage of the polylinker region at the first and second restriction sites, following conversion of the circular single-stranded recombinant vector molecule into a circular double-stranded vector molecule, will produce a linearized double-stranded vector molecule having a first terminus adjacent the cloning region that is susceptible to digestion by Exonuclease III and a second terminus, not susceptible to digestion by Exonuclease III, adjacent the sequencing primer binding site.

3. The reagent kit of Claim 1, wherein the DNA polymerase is selected from among T_4 DNA polymerase and *E. coli* DNA polymerase III holoenzyme.

4. The reagent kit of Claim 3, wherein the DNA polymerase is T_4 DNA polymerase.

5. The reagent kit of Claim 1, further comprising a DNA ligase.

6. The reagent kit of Claim 5, wherein the DNA ligase is T_4 DNA ligase.

7. The reagent kit of Claim 1, further comprising T_4 polynucleotide kinase.

8. The reagent kit of Claim 3, further comprising T_4 DNA polymerase.

9. The reagent kit of Claim 1, further comprising the four deoxynucleoside triphosphates.

10. The reagent kit of Claim 1, further comprising one or more solutions selected from the group consisting of DNA polymerase buffer, Exonuclease III digestion buffer, single-strand-specific endonuclease digestion buffer, single-strand-specific endonuclease termination buffer, and DNA ligase buffer.

11. The reagent kit of Claim 1, further comprising a description of a process for producing shortened DNA fragments of a target DNA segment, the process comprising the steps of: inserting a target DNA segment into the single-stranded vector molecule; cloning the single-stranded vector molecule containing the target DNA segment; annealing the first polymerase primer to the cloned single-stranded vector molecules; extending the first polymerase primer with the DNA polymerase to produce a plurality of circular double-stranded vector molecules with a nick in the newly synthesized (-) strand 5' adjacent to the first polymerase primer sequence; unidirectionally digesting the nicked strand of the circular double-stranded vector molecules from the 3' end of the nick with the Exonuclease III; removing, at timed intervals during the unidirectional digestion, a portion of the Exonuclease III-gapped circular vector molecules; removing single-stranded regions of the Exonuclease III-gapped circular vector molecules with the single-strand-specific endonuclease, to produce shortened linear double-stranded DNA fragments comprising all of the vector DNA; recircularizing the shortened DNA molecules; and cloning the shortened DNA molecules to obtain a plurality of circular DNA molecules containing overlapping target DNA fragments derived from the target DNA segment.

12. The reagent kit of Claim 1, further comprising a description of a process for producing shortened DNA fragments of a target DNA segment, the process comprising the steps of: inserting a target DNA segment into the single-

stranded vector molecule; cloning the single-stranded vector molecule containing the target DNA segment; annealing a custom polymerase primer to the target DNA segment in the cloned single-stranded vector molecules; extending the custom polymerase primer with the DNA polymerase to produce a plurality of circular double-stranded vector molecules with a nick in the newly synthesized (-) strand 5' adjacent to the custom polymerase primer sequence in the target DNA segment; unidirectionally digesting the nicked strand of the circular double-stranded vector molecules from the 3' end of the nick with the Exonuclease III; removing, at timed intervals during the unidirectional digestion, a portion of the Exonuclease III-gapped circular vector molecules; removing single-stranded regions of the Exonuclease III-gapped circular vector molecules with the single-strand-specific endonuclease, to produce shortened linear double-stranded DNA fragments comprising all of the vector DNA; recircularizing the shortened DNA molecules; and cloning the shortened DNA molecules to obtain a plurality of circular DNA molecules containing overlapping target DNA fragments derived from the target DNA segment.

13. A process for producing shortened DNA fragments of a target DNA segment, comprising:

producing a plurality of circular single-stranded vector molecules, each vector molecule (+ strand) comprising vector DNA comprising a polylinker region, the polylinker region comprising a first polymerase primer binding site 5' to a cloning region, the cloning region comprising the target DNA segment;

annealing a first polymerase primer to the first polymerase primer binding site of the circular single-stranded vector molecules;

extending the first polymerase primer with DNA polymerase through first the vector DNA and then the target DNA segment up to the 5' side of the first polymerase primer, to produce a plurality of circular double-stranded vector molecules with a nick in the newly synthesized (-) strand 5' adjacent to the first polymerase primer sequence;

unidirectionally digesting the nicked strand of the circular double-stranded vector molecules from the 3' end of the nick with Exonuclease III;

removing, at timed intervals during the unidirectional digestion, a portion of the Exonuclease III-gapped circular vector molecules;

removing single-stranded regions of the Exonuclease III-gapped circular vector molecules, to produce shortened linear double-stranded DNA fragments comprising all of the vector DNA;

recircularizing the shortened DNA molecules; and

cloning the shortened DNA molecules to obtain a plurality of circular DNA molecules containing overlapping target DNA fragments derived from the target DNA segment.

14. A process for producing shortened DNA fragments of a target DNA segment, comprising:

producing a plurality of circular single-stranded vector molecules, each vector molecule (+ strand) comprising vector DNA comprising a polylinker region, the polylinker region comprising the target DNA segment;

annealing a custom polymerase primer to the target DNA segment in the circular single-stranded vector molecules;

extending the custom polymerase primer with DNA polymerase through the circular single-stranded vector molecules up to the 5' side of the custom polymerase primer, to produce a plurality of circular double-stranded vector molecules with a nick in the newly synthesized (-) strand 5' adjacent to the custom polymerase primer sequence in the target DNA segment;

unidirectionally digesting the nicked strand of the circular double-stranded vector molecules from the 3' end of the nick with Exonuclease III;

removing, at timed intervals during the unidirectional digestion, a portion of the Exonuclease III-gapped circular vector molecules;

removing single-stranded regions of the Exonuclease III-gapped circular vector molecules, to produce shortened linear double-stranded DNA fragments comprising all of the vector DNA;

recircularizing the shortened DNA molecules; and

cloning the shortened DNA molecules to obtain a plurality of circular DNA molecules containing overlapping target DNA fragments derived from the target DNA segment.

15. A reagent kit for producing shortened DNA fragments of a target DNA segment, comprising in combination:

at least one polymerase primer;

a DNA polymerase capable of extending the polymerase primer;

Exonuclease III;

a single-strand-specific endonuclease; and

a description of a process comprising the steps of: inserting a target DNA segment into a single-stranded vector molecule; cloning the single-stranded

vector molecule containing the target DNA segment; annealing the p lymerase primer to the cloned single-stranded vector molecules; extending the polymerase primer with the DNA polymerase to produce a plurality of circular double-stranded vector molecules with a nick in the newly synthesized (-) strand 5' adjacent to the polymerase primer sequence; unidirectionally digesting the nicked strand of the circular double-stranded vector molecules from the 3' end of the nick with the Exonuclease III; removing, at timed intervals during the unidirectional digestion, a portion of the Exonuclease III-gapped circular vector molecules; removing single-stranded regions of the Exonuclease III-gapped circular vector molecules with the single-strand-specific endonuclease, to produce shortened linear double-stranded DNA fragments comprising all of the vector DNA; recircularizing the shortened DNA molecules; and cloning the shortened DNA molecules to obtain a plurality of circular DNA molecules containing overlapping target DNA fragments derived from the target DNA segment.

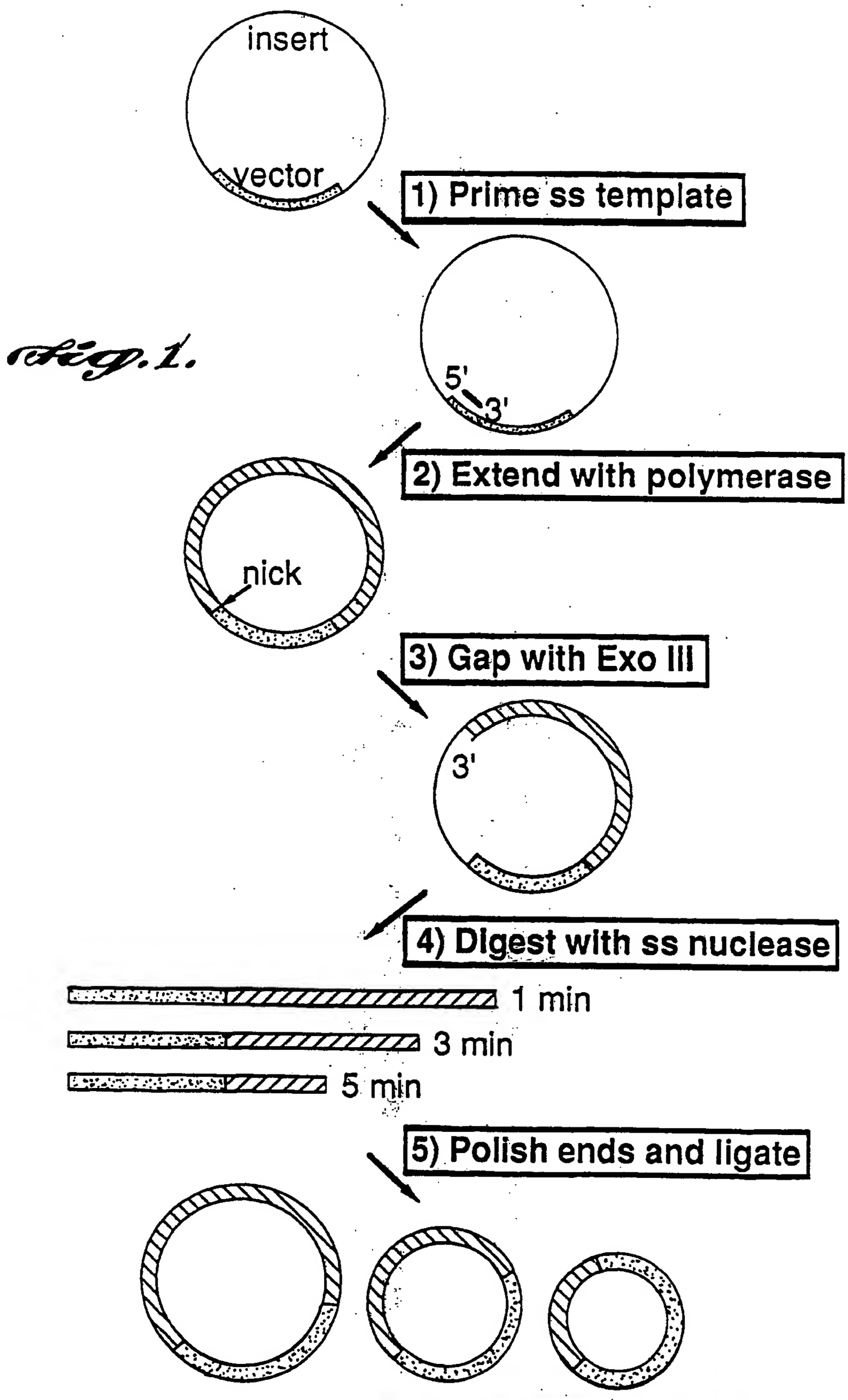
insert
vector
1) Prime ss template
5'
3'
2) Extend with polymerase
nick
3) Gap with Exo III
3'
4) Digest with ss nuclease
1 min
3 min
5 min
5) Polish ends and ligate
6) Transform E. coli

Fig. 1.

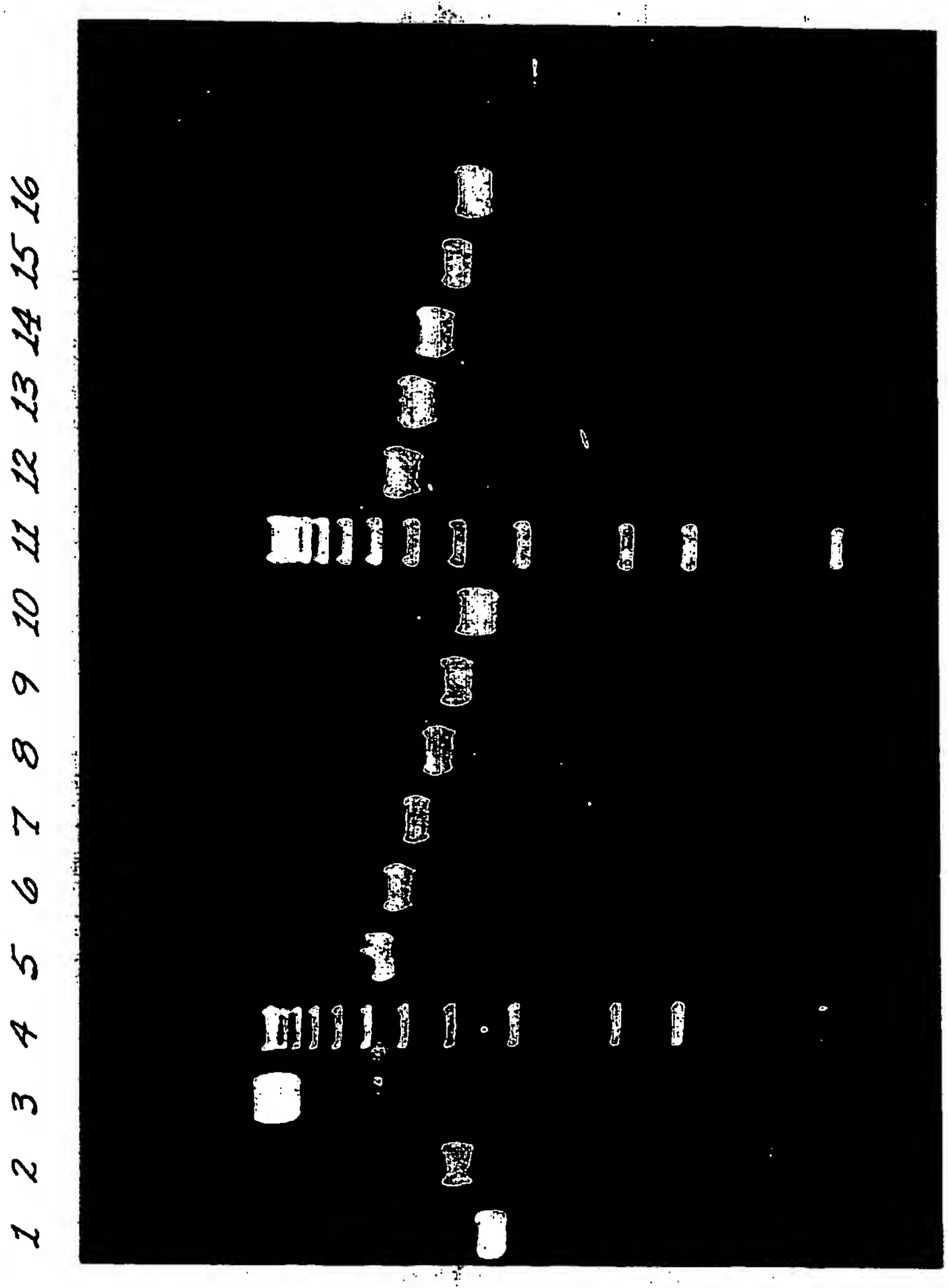

Fig. 2.

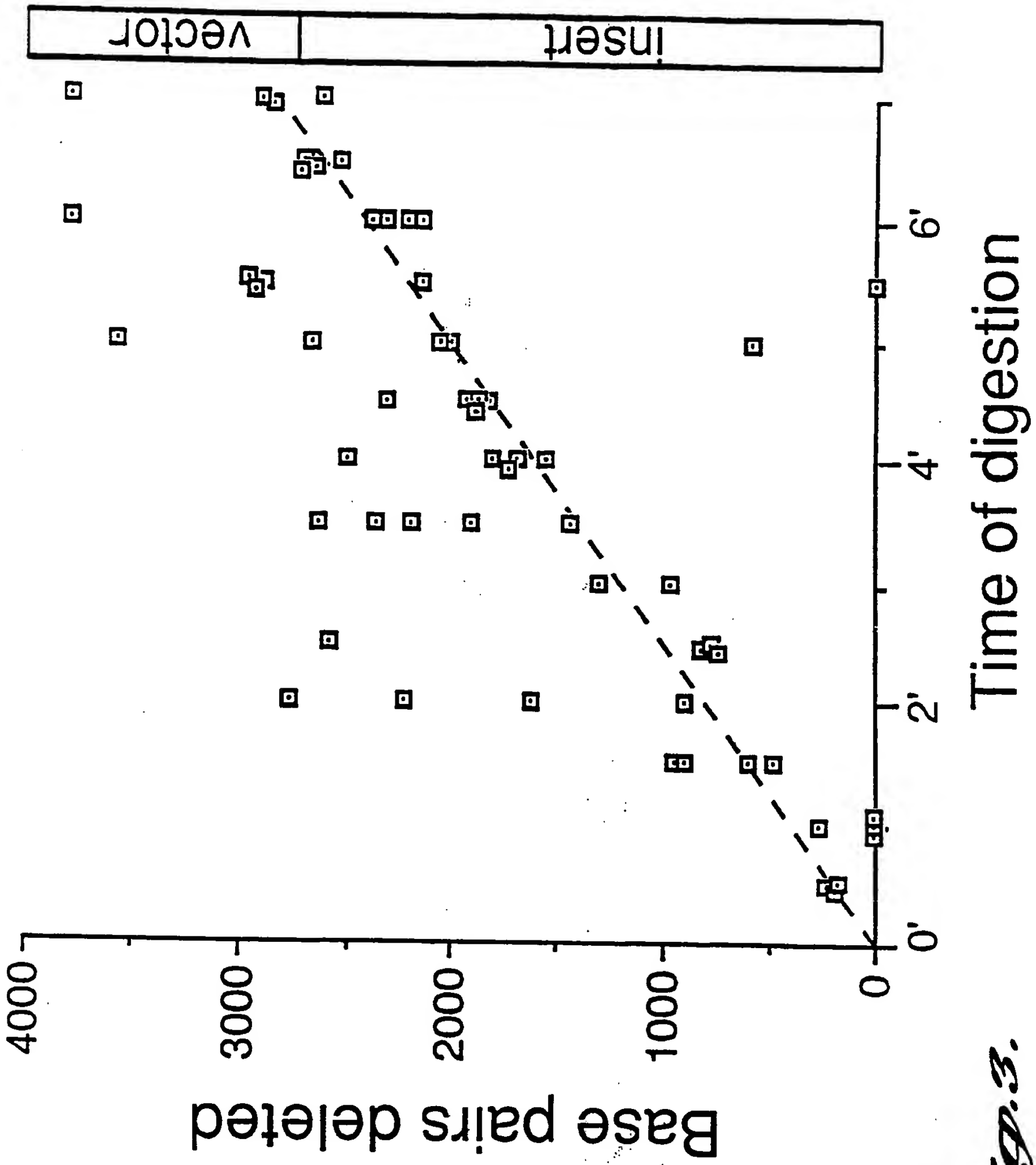
Vector
Insert
Base pairs deleted
4000
3000
2000
1000
0
0'
2'
4'
6'
Time of digestion

Fig. 3.

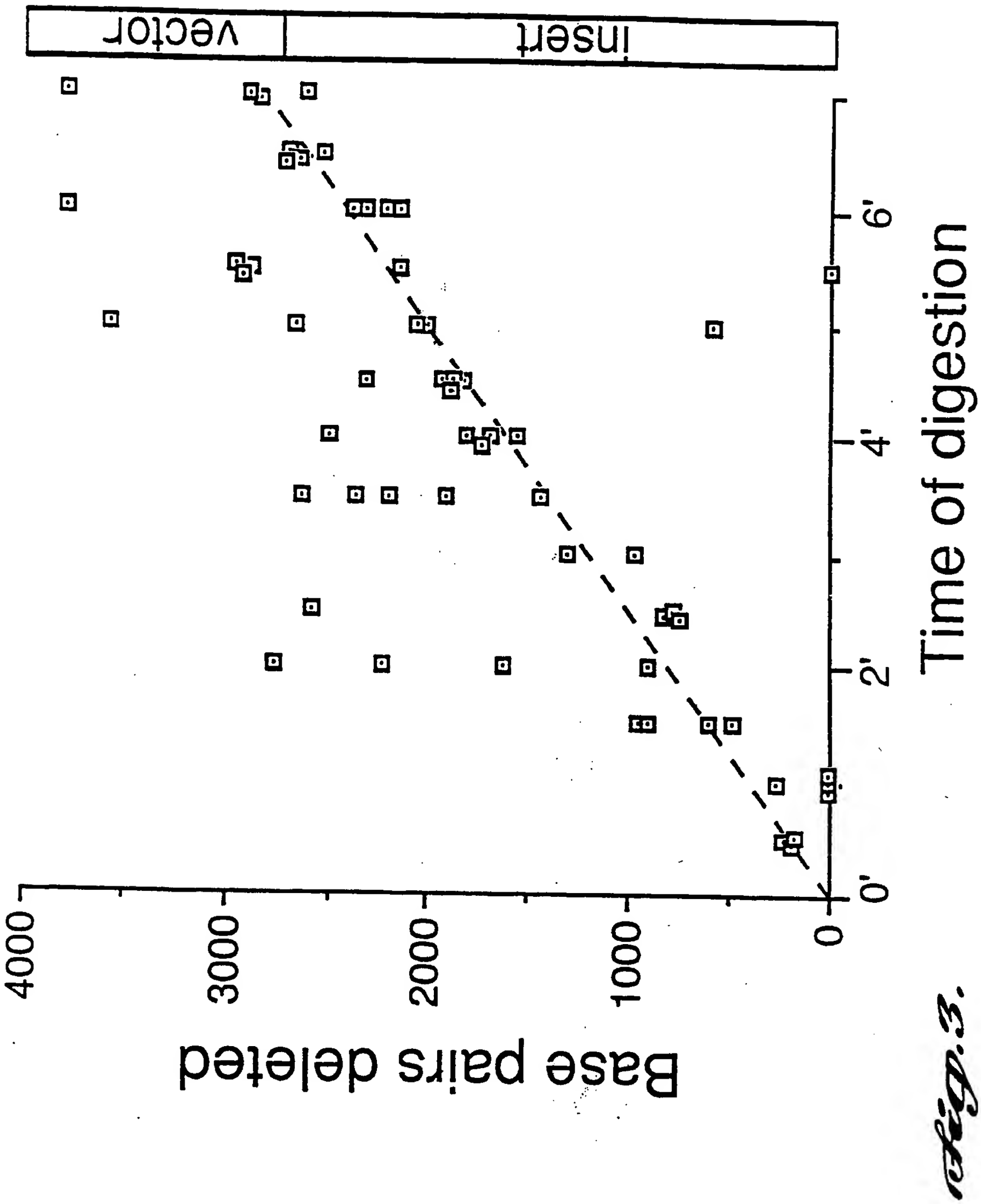
Vector
Insert
Time of digestion
0'
2'
4'
6'
Base pairs deleted
0
1000
2000
3000
4000

Fig. 3.

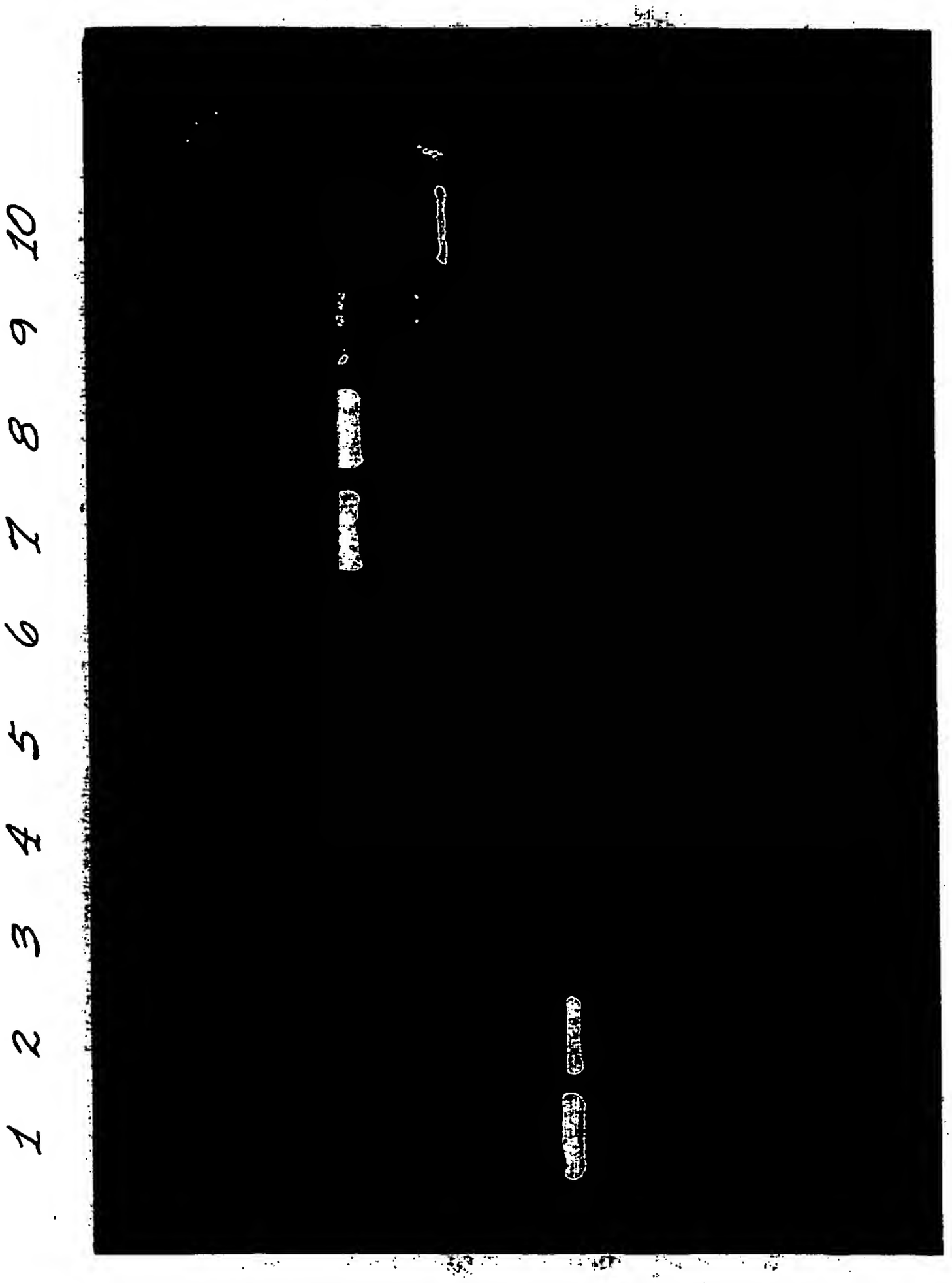

Fig. 4.

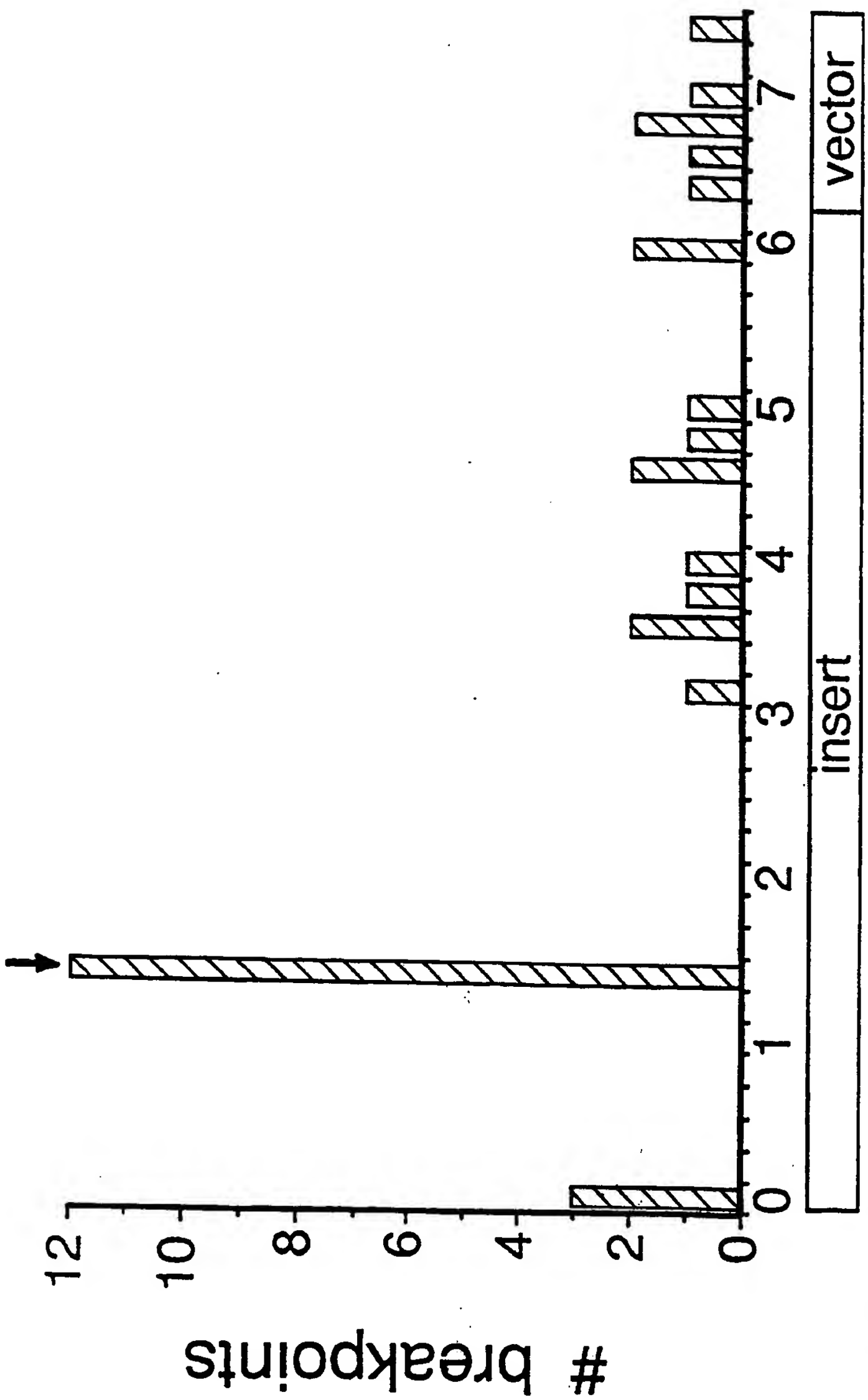
vector
insert
kilobases deleted
0
1
2
3
4
5
6
7
breakpoints
0
2
4
6
8
10
12

Fig. 5.

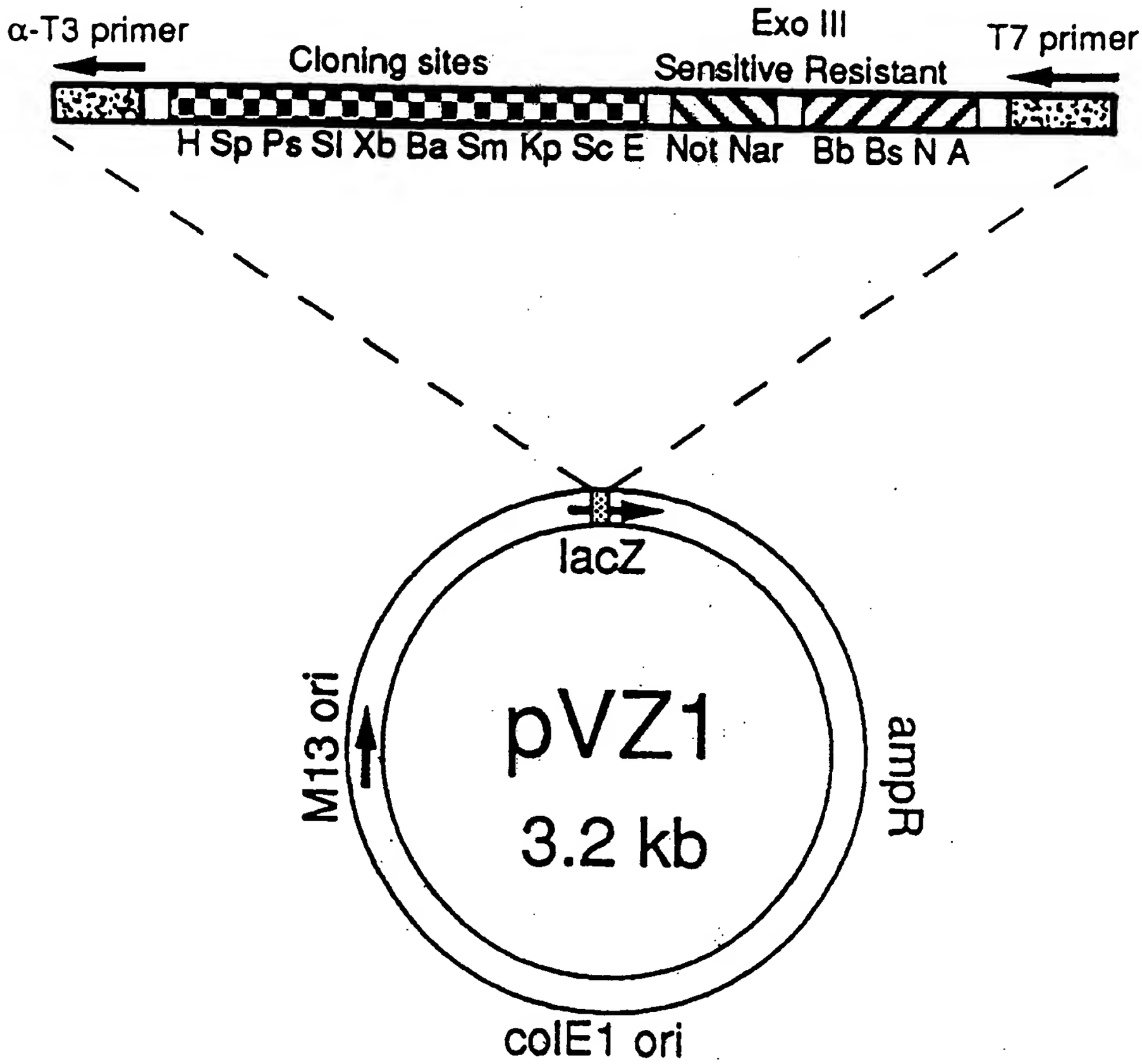
α-T3 primer
Cloning sites
Exo III
Sensitive Resistant
T7 primer
H Sp Ps Sl Xb Ba Sm Kp Sc E Not Nar Bb Bs N A
lacZ
M13 ori
pVZ1
3.2 kb
ampR
colE1 ori

Fig. 6a.

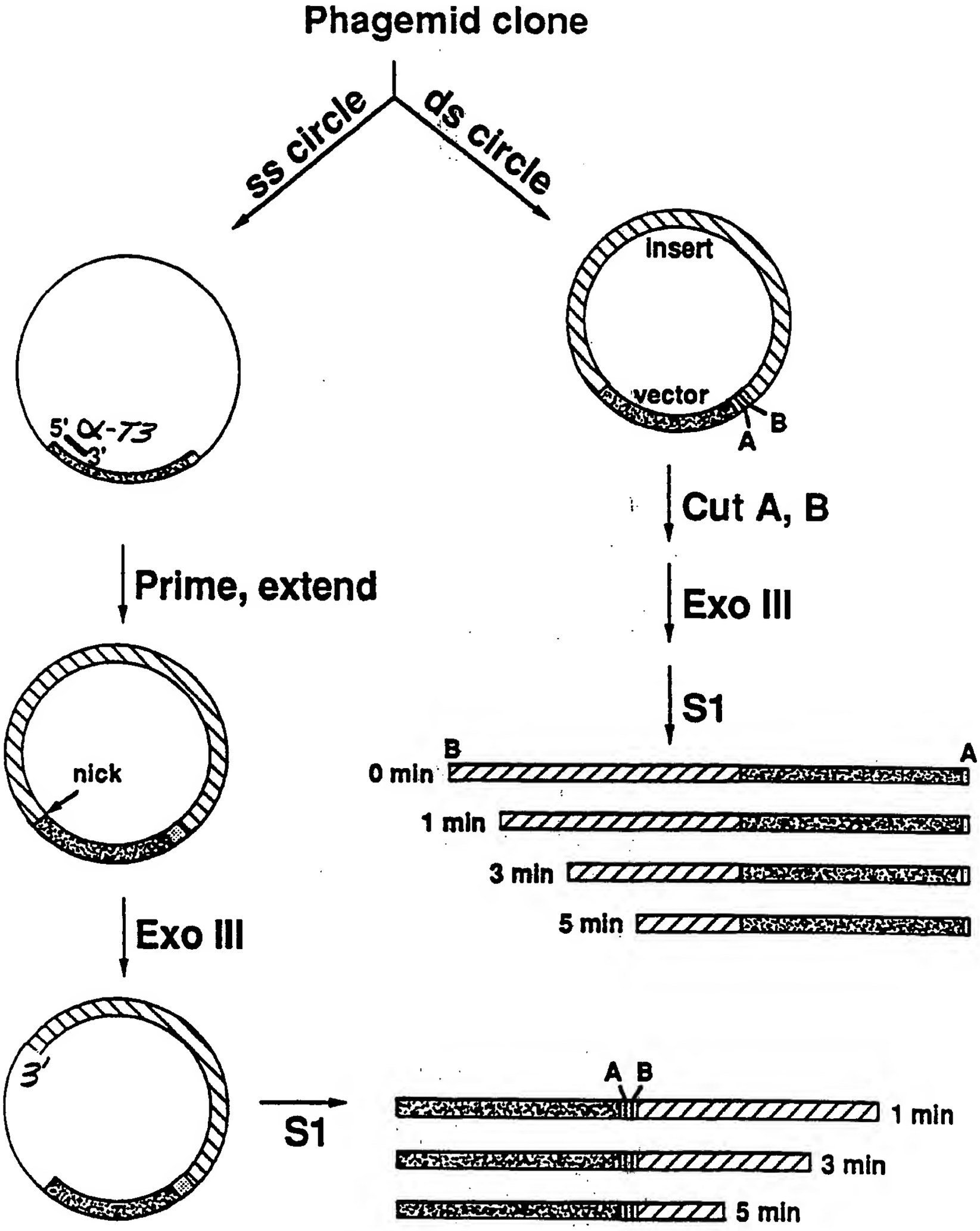
Phagemid clone
ss circle
ds circle
5' α-T3 3'
Prime, extend
nick
Exo III
3'
S1
insert
vector
A
B
Cut A, B
Exo III
S1
B
A
0 min
1 min
3 min
5 min
A B
1 min
3 min
5 min

Fig. 6b.

INTERNATIONAL SEARCH REPORT

International Application No. PCT/US91/ 00951

I. CLASSIFICATION SUBJECT MATTER (if several classification symbols apply, indicate all) [6]

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC(5): C12Q 1/68, 1/42; C12N 9/12, 15/11, 15/66; C12P 19/34.
U.S. CL: 435/6, 21,172.3, 194, 320.1; 935/77.

II. FIELDS SEARCHED

Minimum Documentation Searched [7]

Classification System	Classification Symbols
U.S.	435/6, 21, 172.3, 194, 320.1; 935/77

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched [8]

DATABASES. Dialog Online Services; Automated Patent System (File USPAT, 1975-1991). Keywords: Exonuclease, Delet? Gap, Nick

III. DOCUMENTS CONSIDERED TO BE RELEVANT [9]

Category [10]	Citation of Document, [11] with indication, where appropriate, of the relevant passages [12]	Relevant to Claim No. [13]
Y	NUCLEIC ACIDS RESEARCH, Vol. 11, No.2, Issued 1983, BARNES et al., "Kilo-sequencing: An Ordered Strategy for Rapid DNA Sequence Data Acquistion," pages 349-368, see e.g., Figure 3.	1-15
Y	BIOTECHNIQUES, Vol. 7, No. 7. issued July 1989, LIU et al., "Rapid Generation of Subclones for DNA Sequencing Using the Reverse Cloning Procedure," pages 722-728, see e.g., Figure 1.	1-15
Y	METHODS IN ENZYMOLOGY, Vol. 65, issued 1980, ROGERS et al., "Exonuclease III of Escherichia coli K-12, an AP Endonuclease," pages 201-211, see entire document.	1-15

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search	Date of Mailing of this International Search Report
17 May 1991	09 JUL 1991
International Searching Authority	Signature of Authorized Officer
ISA/US	Richard Lebovitz

Form PCT/ISA/210 (second sheet) (Rev. 11-87)

PCT US 91/00951

International Application No.

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)

Category *	Citation of Document, [16] with indication, where appropriate, of the relevant passages [17]	Relevant to Claim No [18]
Y	PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCE, Vol. 79, issued March 1982, SHORTLE et al., "Gap Misrepair Mutagenesis: Efficient Site-Directed Induction of Transition, Transversion, and Frame-Shift Mutations In vitro," pages 1588-1592, see entire document.	1-15